BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 5: e74141
doi: 10.3897/biss.5.74141

OPEN ACCESS

Conference Abstract

#RetroPIDs: The missing link to the foundation of biodiversity knowledge

Nicole Kearney[‡,§,|], Colleen Funkhouser[¶,#], Mike Lichtenberg[¶], Bess Missell[#], Roderic Page[¤], Joel Richard[#], Diane Rielinger[«], Susan Lynch[»]

‡ Biodiversity Heritage Library Australia, Melbourne, Australia
§ Atlas of Living Australia, Melbourne, Australia
| Museums Victoria, Melbourne, Australia
¶ Biodiversity Heritage Library, Washington D.C., United States of America
# Smithsonian Libraries and Archives, Smithsonian Institution, Washington D.C., United States of America
¤ University of Glasgow, Glasgow, United Kingdom
« Botany Libraries, Harvard University Herbaria, Cambridge, United States of America
» The New York Botanical Garden, New York, United States of America

Corresponding author: Nicole Kearney (nkearney@museum.vic.gov.au)

Received: 08 Sep 2021 | Published: 08 Sep 2021

Citation: Kearney N, Funkhouser C, Lichtenberg M, Missell B, Page R, Richard J, Rielinger D, Lynch S (2021) #RetroPIDs: The missing link to the foundation of biodiversity knowledge. Biodiversity Information Science and Standards 5: e74141. https://doi.org/10.3897/biss.5.74141

Abstract

The Biodiversity Heritage Library (BHL) will soon upload its 60 millionth page of open access biodiversity literature onto the BHL website and the BHL's Internet Archive Collection. The BHL's massive repository of free knowledge includes content that is available nowhere else online, as well as accessible versions of content that are locked behind paywalls elsewhere. If we are to continue to expand our understanding of life on Earth, we must ensure that the foundation of biodiversity knowledge provided by BHL is discoverable by the tools we rely on to navigate the ever-expanding internet. These tools – search engines and their algorithms – preferentially deliver (and rank) content with good metadata and persistent identifiers (PIDs).

In modern online publishing, PID assignment and linking happens at the point of publication: DOIs (Digital Object Identifiers) for publications, ORCIDs (Open Researcher and Contributor IDs) for people, and RORs (Research Organization Registry IDs) for organisations. The DOI system provided by Crossref (the DOI registration agency for scholarly content) delivers reciprocal citations, enabling convenient clicking from article to

This is an open access article distributed under the terms of the CC0 Public Domain Dedication.

article, and citation tracking, enabling authors and institutions to track the impact and reach of their research output. Publications that lack PIDs, which include the vast majority of legacy literature, are hard to find and sit outside the linked network of scholarly research. This makes it nearly impossible to determine whether they are being cited, let alone viewed, mentioned, shared or liked.

At TDWG 2020, Page 2020, Kearney 2020, Richard 2020 (and 2019, Page 2019b, Page 2019a, Kearney 2019b, Kearney 2019a and 2018, Kearney 2018), we emphasised the need to bring the historic biodiversity literature into the modern linked network of scholarly research. In October 2020, BHL launched a new working group to do exactly this. The BHL Persistent Identifier Working Group (Team #RetroPID) brings together expertise from across BHL's global community. Over the past year, we have worked tirelessly to make it easier to find, cite, link, share and track the content on BHL, adding article-level metadata to journals and retrospectively assigning DOIs (#RetroPIDs). Most importantly, we have developed the tools and documentation that will enable the entire BHL community to take contributed content from "just" accessible to persistently discoverable.

This paper will detail our efforts to retrofit the historic literature (a square peg) into the modern PID system (a round hole) and will present both the achievements and the challenges of this important work.

Keywords

persistent identifiers, PIDs, digital object identifiers, DOIs, Biodiversity Heritage Library, BHL, metadata, publishing, research, open access, paywalls, online content, Crossref, Unpaywall, citations, biodiversity knowledge graph, linked data, ISSNs, copyright, literature, accessibility, discoverability, FAIR

Presenting author

Nicole Kearney

Presented at

TDWG 2021

References

- Kearney N (2018) What are we DOIng about the Missing Links? Connecting Taxonomic Names to the Linked Network of Scholarly Research. Biodiversity Information Science and Standards 2 https://doi.org/10.3897/biss.2.25700

- Kearney N (2019a) It's Not Always FAIR: Choosing the Best Platform for Your Biodiversity Heritage Literature. Biodiversity Information Science and Standards 3 https://doi.org/10.3897/biss.3.35493
- Kearney N (2019b) If Wikipedia is the Gateway to Biodiversity Knowledge, How do we Open the Gate? Biodiversity Information Science and Standards 3 https://doi.org/10.3897/biss.3.37356
- Kearney N (2020) Discovering the Platypus: From its scientific description to its DOI. Biodiversity Information Science and Standards 4 https://doi.org/10.3897/biss.4.59089
- Page R (2019a) Text-mining BHL: towards new interfaces to the biodiversity literature. Biodiversity Information Science and Standards 3 https://doi.org/10.3897/biss.3.35013
- Page R (2019b) Wikidata and the biodiversity knowledge graph. Biodiversity Information Science and Standards 3 https://doi.org/10.3897/biss.3.34742
- Page R (2020) Strategies for Assembling the Biodiversity Knowledge Graph. Biodiversity Information Science and Standards 4 https://doi.org/10.3897/biss.4.59126
- Richard J (2020) Improving Taxonomic Name Finding in the Biodiversity Heritage Library. Biodiversity Information Science and Standards 4 https://doi.org/10.3897/biss.4.58482